Nonparametric Bayes Methods for Big Data in Neuroscience

Duke University

PI: John Pearson

Grant Number: 1-K01S25442-01

I am applying for mentored career development through the BD2K initiative to gain the skills and expertise necessary to transition to an independent research career developing methods for the analysis of "big data" in systems and cognitive neuroscience. Following my Ph.D. training in theoretical physics, I transitioned into computational neuroscience, where I have focused on problems in the neurophysiology of reward and decision-making, particularly models of reinforcement learning and choice behavior. For the last five years, I have also gained extensive experience in electrophysiological recording in both human surgical patients and non-human primates, deepening my appreciation of the difficulties involved in analyzing real neuroscience data. During this time, I have become convinced that the single most pressing challenge for neuroscience in the next decade will be the problem of how we process, analyze, and synthesize the rapidly expanding volumes of data made available by new technologies, and as I transition to the faculty level, I am seeking to orient my own research program toward these goals. To do so, I will need to complement my strong quantitative background and electrophysiological recording skills with specific training in machine learning, signal processing, and analysis of data from functional magnetic resonance imaging (fMRI). I am focusing on the first because the statistics of data analysis are an essential core competency for any big data researcher; on the second because understanding the methods by which we process and acquire data are as essential as how we analyze them; and on the third because not only are fMRI data among the most readily available large datasets, but effective analysis of fMRI data will have immediate clinical applications. For this project, I have assembled a team of mentors with strong and overlapping expertise in these three areas. These mentors have committed to support my transition to a focus on big data research, an approach that builds on multiple existing collaborations I have with laboratories at Duke. My ultimate goal is to head a lab in which I apply the skills and training I acquire during the award period to developing computational methods that will harness the power of big data to answer fundamental questions in cognitive and translational neuroscience. Environment. Duke University is home to outstanding resources in both neuroscience and big data research. Its interdisciplinary big data effort, the Information Initiative at Duke, brings together researchers from statistics, computer science, and electrical engineering with those in genetics, neuroscience, and social science to facilitate collaboration across the disciplines. The Duke Institute for Brain Sciences, with which I am affiliated, comprises over 150 faculty across the brain sciences at Duke, from clinicians to biomedical engineers. I will be mentored by Dr. David Dunson, a recognized leader in Bayesian statistical methods for machine learning, along with Dr. Lawrence Carin and Dr. Guillermo Sapiro, experts in signal and image processing and machine learning and frequent collaborators with Dr. Dunson. In addition Dr. Scott Huettel, an expert in fMRI and author of a leading neuroimaging textbook, will oversee my training in fMRI data analysis. Moreover, I will have access to data from a large and diverse pool of laboratories at Duke, including one of the largest neuroimaging datasets in the country. Most importantly, Duke is fully committed to supporting me with the resources and time necessary to pursue the training outlined in this career development award. Research. Each year, one in four adults suffers from a diagnosable mental disorder, with 1 in 25 suffering from a serious mental illness. Yet our ability to anticipate the onset of mental illness - even our ability to understand its effets within the brain - has been limited by the recognition that these diseases are not primarily disorders of independent units, but patterns of pathological brain activation. However, we currently lack a meaningful characterization of patterns of activity within neural networks, and thus the ability to discuss, discover, and treat them effectively. Yet an improvement in our abilit to characterize and detect these patterns would result in major clinical impact. Therefore, under the guidance of my mentoring team, I propose to characterize patterns of network activity in neuroscience datasets using methods from machine learning. Because many mental illnesses are typified either by a pathological relationship between sufferers and stimuli in the world (post traumatic stress disorder, eating disorders) or intrinsic patterns of disordered thought (major depression, obsessive-compulsive disorder), I focus on three key questions for pattern detection: 1) How does the brain encode complex, unstructured stimuli? 2) What are the basic building blocks of healthy and diseased patterns of intrinsic brain activity? 3) How do patterns of brain activity change in response to changes in behavioral state? My approach makes use of recent advances in Bayesian nonparametric methods, as well as fast variational inference approaches that scale well to large datasets. In addition, because the datasets I will use, fMRI and electrophysiology data, are particular examples of the much larger class of multichannel time series data, the results will apply more broadly to other types of data, in neuroscience and beyond. PUBLIC HEALTH RELEVANCE: Recent evidence suggests that most mental illnesses result from disruptions in normal patterns of brain activity. Yet discovering, detecting, and describing these patterns of activity has proven difficult to do in conventional experiments. This project wil develop computer algorithms for pattern detection that can be applied to the scientific study and early diagnosis of mental illness.